DOCUMENT RESUME

ED 337 491 TM 017 327

TITLE Using the School Report.
INSTITUTION Massachusetts State Dept. of Education, Boston.
PUB DATE Nov 90
NOTE 42p.; Some pages contain extremely small print. For a related report, see TM 017 328.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Educational Assessment; *Evaluation Utilization; Intermediate Grades; *Profiles; *Program Evaluation; Scores; Secondary Education; Standardized Tests; *State Programs; Student Evaluation; *Testing Programs; Test Results; Test Use
IDENTIFIERS *Massachusetts Educational Assessment Program

ABSTRACT

Schools that are concerned with evaluating their programs need more than just the average scaled assessment scores from standardized tests. To help schools, the Massachusetts Educational Assessment Program produces school and district reports that contain the stories behind the averages. These reports describe the kind of students who took the tests (given in grades 4, 8, and 12), present performance profiles of different groups of students, and provide detailed analyses of performance. This booklet is designed to help schools use their reports to the fullest. Chapter 1 introduces the assessment program. Chapter 2 provides background information and reviews the intent of the program and how it has been implemented. In Chapter 3, the kinds of information in the reports are reviewed, with demonstrations of the kinds of comparisons that can be made. Chapter 4, reproduced from previous editions of this booklet, provides suggestions and examples of how to use the results. Appendix I is a short description/definition of the skills and knowledge of students at various proficiency levels. Appendix II provides tables of the mean and standard deviations of the raw scores for each content subcategory. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

TM

CHAPTER 188

Massachusetts Educational Assessment Program

September 1990

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. RAYNOLDS, JR.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED337491

Using the School Report

MASSACHUSETTS DEPARTMENT OF EDUCATION

BEST COPY AVAILABLE

TM017327

MASSACHUSETTS BOARD OF EDUCATION

Mr. James F. Crain, *Chairperson*, Gloucester
Dr. Raquel Bauman, *Vice Chairperson*, Holden

Mr. John J. Gould, Boston
Mr. William K. Irwin, Jr., Wilmington
Mr. Joseph E. Killory, Sandwich
Dr. Anne S. Larkin, Bedford
Dr. Richard R. Rowe, Belmont
Mr. Jonathan D. Silverman, Worcester
Ms. Theodora A. Sylvester, Springield
Mrs. Frances M. Turner, South Egremeont
Dr. Joan Wallace-Benjamin, Boston
Mr. Frederick A. Wang, Needham

Dr. Randolph W. Bromery, *ex officio Board Member, Interim Chancellor, Board of Regents of Higher Education*

Harold Raynolds, Jr., *Commissioner of Education, Secretary of the Board of Education*

OFFICE OF PLANNING, RESEARCH AND EVALUATION

Allan S. Hartman, Director of Research

M. Elizabeth Badger
Beverly Miyares
Brenda Thomas
Mary Robertson

The Massachusetts Department of Education insures equal employment/educational opportunities/affirmative action regardless of race, color, creed, national origin or sex, in compliance with Title IV and Title IX, or handicap, in compliance with section 504.

USING THE SCHOOL REPORT

Massachusetts Educational Assessment Program

MASSACHUSETTS DEPARTMENT OF EDUCATION
November 1990

TABLE OF CONTENTS

Chapter 1
Introduction 1

Chapter 2
Background: The Massachusetts Educational Assessment Program 3

Chapter 3
Making Comparisons 5

Chapter 4
Developing an Action Plan 14

Appendix I 22

Appendix II 29

CHAPTER ONE

INTRODUCTION

One of my feet is submerged in boiling water; the other is frozen into an iceblock. On the average, I'm doing just fine.

Anyone with some knowledge of statistics knows how incomplete an average can be. When reported as an achievement score, it tells nothing about the range of individual scores or how they are distributed. More important for schools, however, it tells nothing about the circumstances behind the scores. Schools with the same average scores can differ dramatically in regard to their students' performance. In one school, the students may be extremely homogeneous; in another, a small group of very high achievers may be balanced by a much larger group of low-achieving students. Yet, both may have the same score. Schools with the same average may also differ widely in terms of their curriculum. A school that treats all topics relatively superficially, attending only to the main ideas, may score the same as one in which some topics are studied in depth while others are ignored. The schools may both attain the same average on the topic as a whole.

The average scaled assessment scores which are reported in your school report have these limitations. Schools that are concerned with evaluating their programs will not be satisfied with these numbers alone. In order to help them, the Massachusetts Educational Assessment Program produces school and district reports that contain the stories behind the averages. These reports describe the types of students who took the test. They give a performance profile of different groups of students. They give a detailed analysis of performance on different topics within each subject area. They paint a picture of how students perform on more demanding open-ended questions, to what extent they use their time to reinforce learning, and the extent to which they have studied different topics.

The purpose of this booklet is to help schools use these reports to their fullest. Chapter Two provides background on the assessment program, discussing its intent and how that purpose has been carried out. In Chapter Three, we review the various types of information in the report, demonstrating how kinds of comparisons that can be made. Chapter 4 is reproduced from the previous edition of this publication, and provides suggestions and examples of how to use the results. We have also included two appendices that you may find useful in interpreting the reports. Appendix I gives short descriptions of the skills and knowledge of students at various proficiency levels. (More complete descriptions will be found in the publication. *Description of the MEAP*

Proficiency Scales.) Appendix II provides the mean and standard deviation of the raw scores for each content sub-category.

The purpose of the assessment is to provide information to schools—information that is often costly or impossible to obtain within the system itself. The process should not stop there, however. In a certain sense, the assessment reports which you have received should represent only a beginning. They should be used, not to give answers but to generate questions. We hope that this handbook will help you understand the kinds of questions that can be asked and the ways in which answers can be found.

CHAPTER TWO

BACKGROUND: THE MASSACHUSETTS EDUCATIONAL ASSESSMENT PROGRAM

The Massachusetts Educational Assessment Program was established in 1985 as one of a number of programs designed to improve curriculum and instruction in the public schools in Massachusetts. Under the mandate by which it was established, it is designed with two purposes:

- to provide comparisons of performance at school, district, state, and national levels, when possible;
- to improve curriculum and instruction.

In 1985-86, the first year of the assessment, third, seventh, and eleventh grade students were tested in reading, mathematics, and science. In the spring of 1988 the grades at which students were tested were changed to fourth, eighth, and twelfth grades, and social studies was added to the subject areas tested.

Test Development

The objectives which form the framework for each content area are described in a separate publication, *Description of Test Content and Reporting Categories*. These objectives were arrived at through an iterative process that involved national curriculum experts, state curriculum experts, the Department of Education, and the schools themselves.

Initially, the Curriculum Advisory Committees, composed of teachers and curriculum coordinators throughout the state, reviewed the set of objectives drawn up by the National Assessment of Educational Progress, as well as those from other sources. The committees used this material to construct a tentative objectives framework for each subject which included the content categories and types of thinking that are considered to be the essential components for achievement within that discipline They also drew up a list of content topics for consideration by schools. All schools were asked to react to this list by rating each topic in terms of its importance for student learning by the target grade level (4, 8, and 12). The Curriculum Advisory Committees then used the results of this survey to finalize the objectives framework and to decide what specific aspects of knowledge or skills should be included within each specific subcategory of the framework.

These frameworks are not static. During the last few years, educators in various subject areas have been reviewing and redefining their understanding of how learning takes

place, as well as their definition of what it means to know a subject. The Department of Education has attempted to respond to these ideas while continuing to maintain the continuity of the Assessment Program. As a result, the Mathematics Framework has been modified to reflect the emphasis placed by the National Council of Teachers of Mathematics on problem solving and reasoning. Likewise, the Reading Framework has been expanded to capture the role that reading strategies play in constructing meaning.

Components of the Assessment

Since the assessment was established in 1985, a movement has developed for standardized testing to reflect more clearly the kinds of behavior that are considered to be the goals of education. As a result, questions that call for application, analysis, and evaluation have formed a larger part of the multiple-choice section of the assessment. In addition, one test form at each grade level contains a set of open-ended questions which demand the kinds of thinking that cannot be captured by an item that asks only for the identification of a correct option. These open-ended questions primarily focus on students' beliefs and understanding, as well as upon their ability to apply knowledge in less traditional contexts. Furthermore, in 1989, performance testing was added to the program. A sample of fourth and eighth grade students were tested on their ability to apply mathematical and scientific concepts in solving problems. The results of these tests were reported in a separate series of brochures entitled *Beyond Paper and Pencil*.[1]

The three testing formats—multiple-choice, open-ended, and performance—are designed to provide a rounded picture of what students know, how they apply that knowledge, and how they assimilate knowledge to carry out multi-dimensional tasks.

Other Information

In addition to the cognitive tests, a series of student, teacher, and principal questionnaires have also been developed. Many of the questions are identical to those used by the National Assessment of Educational Progress and other national programs; others were developed by the Department with the advice of the Curriculum Advisory Committees. Responses to the student questionnaires are primarily used to interpret test results. They form the basis for the Comparison Score Bands and are included in the school reports as additional information to schools. Responses to the teacher and principal questionnaires are aggregated and reported at the state level. Their primary use is in the exploration of policy and curriculum issues; however, they are also used as indicators of the overall quality of education in Massachusetts, particularly in terms of national comparisons.

---

[1] The original budget for the 1990 assessment included funds for additional performance testing to take place during 1991. Because of the current statewide financial situation, testing has been cancelled. However, MEAP will continue to support such testing in schools.

CHAPTER THREE

MAKING COMPARISONS

Schools have many ways to chart the progress of individual students: teacher observations, classroom tests, homework assignments, student projects, and standardized tests. However, it is far more difficult for a school or district to step back and chart its own course. When schools ask "How effective is our program? Are we doing the best job possible?", there are few tools at hand for judging their progress.

One way to measure a school's achievement is to compare the total performance scores of its students with those of students in other schools or the state as a whole. Another way is to look at the performance of students in similar schools. Another is to look internally, at performance among groups of students within the school. The school report which you have received gives information on all these ways and more. The purpose of this chapter is to demonstrate the kinds of comparisons you can make and to help you develop a school profile of achievement.

I. Making Comparisons

Comparing one's school to the district or state is an important method for measuring success. However, schools are complex organisms that defy the simplicity of a single numerical scale. In order to understand how well you are achieving your goals vis-a-vis others in the state it is necessary to compare on different levels.

Comparisons with your district and the state.

In order to make meaningful comparisons among schools, across subject areas, and between years, the assessment scores are scaled. That is, the percentage of correct responses to the test questions is transformed to a common measure, which is based on the normal distribution. In 1988, the mean state score was set at 1300, with a standard deviation of 100 points, and a possible range of scores from 1000 to 1600. This meant that 68% of school scores fell between 1200 and 1400; 16% were below 1200, and another 16% were above 1400. Using the same scale and many of the same items, we see that performance has risen modestly since the 1988 assessment with state averages equal to 1310 in grades 4 and 8. The distribution and the range of scores also remain much the same.

On page 3 of your school report you will find the scores which your students achieved in each of the four subject areas tested. Scores are given, not only for your total student body but for three categories of students: those with limited English proficiency (and not in a special needs program), those in special needs programs, and those who are in

neither ("Other"). In reviewing these results, it is important to note the percentage of students in each category within your school who were **not** tested. A marked difference from the state or district average indicates that the results for this category may not be fairly compared with state averages.

SUMMARY OF STUDENTS TESTED BY CATEGORY

Student Category	STATE				DISTRICT				SCHOOL			
	Number Enrolled	Number Included In Report	Percentage of Category	Percentage of Students In Report	Number Enrolled	Number Included In Report	Percentage of Category	Percentage of Students In Report	Number Enrolled	Number Included In Report	Percentage of Category	Percentage of Students In Report
SPECIAL NEEDS	7694	3917	51%	8%	17	9	53%	6%	17	9	53%	6%
BILINGUAL	2094	679	32%	1%	0	0	0%	0%	0	0	0%	0%
OTHER	47231	45554	96%	91%	136	134	99%	94%	136	134	99%	94%
ALL STUDENTS	57019	50150	88%	100%	153	143	93%	100%	153	143	93%	100%

SUMMARY OF TEST RESULTS BY CATEGORY

Student Category	STATE				DISTRICT				SCHOOL			
	Average Scaled Score				Average Scaled Score				Average Scaled Score			
	Reading	Mathematics	Science	Soc. Studies	Reading	Mathematics	Science	Soc. Studies	Reading	Mathematics	Science	Soc. Studies
SPECIAL NEEDS	1050	1080	1190	1140								
BILINGUAL	1020	1160	1110	1090								
OTHER	1340	1330	1320	1320	1450	1380	1330	1420	1450	1380	1330	1420
ALL STUDENTS	1310	1310	1310	1310	1430	1360	1320	1410	1430	1360	1320	1410

Comparisons in terms of program participation.

In your interpretation of scores it is important to be aware of the extent to which these scores are based on the performance of different groups of students. As you know, some students who do not participate fully in the entire school program are tested. Not surprisingly, their scores are usually lower than those of other students. Since schools differ in the proportion of their populations who have special needs or who lack proficiency in English, the relative influence of the scores of these students differs from school to school. Information about the performance of different groups of students can also be found on page 3 of the school report.

When interpreting your results it may be helpful to ask the following questions in regard to program participation:

- How does the percentage of students in special programs (Special Needs or Bilingual) in our school compare with that of the rest of the state and the district?

- To what extent is our total score affected by the performance of students in special programs?

- How does the performance of these students compare with others throughout the state? Based on these results, are we satisfied with our program? When we exclude

these students who fall into these categories from our total score (that is, consider just the scores of those reported as "Other"), how do we compare with other schools in the state or district?

Comparisons in terms of similar schools

On pages 6 through 13 you will also find subject area scores reported in terms of your Comparison Score Band, which will help you compare the scores of your students with those of other students **in schools like yours.** Comparison Score Bands are determined by relating certain student background factors (such as parental education, English fluency, kindergarten attendance) and certain community factors (such as percentage of students receiving free or reduced price lunch and Kind of Community) to performance on the assessment. The relative influence of each factor on performance is determined for the state as a whole, and the resulting formula is used to compute the Comparison Score Band for each individual school. (See page 6 of your report for a comparison of the values of those background factors in your school with those in the state and district.) Since all schools differ with respect to the factors involved, each school has its unique Comparison Score Band. Its width reflects its reliability and is determined in large part by the number of students in the school. Comparison Score Bands are made wide enough so that half the scores in the state lie within the respective Comparison Score Bands.

Background Factor	Average Value		
	State	District	School
Parent Education Index Range: 2 to 8	5.04	5.18	5.18
Free/Reduced Lunch Index Range: 0 to 1	0.87	0.94	0.94
Family Language Index Range: 0 to 1	0.76	0.96	0.96
Kind of Community		5	5

Comparison Score Bands are an important reference point. They signify that factors beyond the control of schools affect students' academic performance. However, it is important to remember that they are **not prescriptive**, but **normative.** In other words, they do not indicate where a school **should** score, only where it might be **expected** to score, given the background factors involved. On the other hand, scores that fall below the band should be examined carefully. They indicate that your students performed less well than other students with similar economic and social backgrounds. Scores that fall outside their Comparison Score Band and that are at least 50 points away from the center of their band should be looked at with extra care. They indicate that your school scored significantly higher or lower than other schools with a similar student body.

In a separate document, the average scores for schools in each Kind of Community (KOC) are given. Schools are grouped into KOCs according to different economic factors in their communities. Unlike Comparison Score Bands, KOCs do not reflect the background factors of the actual students who took the test and are, therefore, grosser measures of similarity. Their advantage lies in the fact that communities within each KOC are named and described whereas the Comparison Score Bands are uniquely determined for each school and therefore, no such similar grouping is possible. On the other hand, a list of schools or districts with Comparison Score Bands similar to yours is available from the Department should you desire it.

Comparisons with previous assessment results.

Pages 6 through 12 also give the results of the 1988 assessment. Since approximately 80% of the items on the tests were identical from one assessment to the next, statistical adjustments were made so that these scores can be directly compared to those of the previous assessment. However, there are several caveats to remember when comparing results from one year to another.

The size of a school is an important consideration. As students within a class can assume a collective personality that is reflected in the work that they produce, the test results for one class can differ dramatically from those of another, even when the teacher or the curriculum remains the same. It is because of this effect that we do not produce scores for schools with fewer than 20 students. For the same reason, schools with fewer than 60 students should view changes in scores with some caution. For such schools, it would be more valuable to look for rising or falling trends since 1986, despite the fact that different grades were tested during the first year of the assessment. Large schools, where the results are based on the responses of several classes, will give a more reliable indication of the average performance in that school than is the case of small schools where the score is dependent upon the performance of only or two classes. Nevertheless, it is usually conceded that scores that go up or down from one occasion to the next are less reliable than scores that show a consistent trend in either direction. A general rule of 50 points can be used as a mark of significant change.

Year	Content Area	Average Test Score			Comparison Score Band	SCHOOL AVERAGE (♦) RELATIVE TO COMPARISON SCORE BAND ( )
		State	District	School		1000 1100 1200 1300 1400 1500 1600
1990	Total Reading	1310	1430	1430	1335 - 1394	
	Vocabulary	1320	1340	1340		
	Literal Comprehension	1300	1500	1500		
	Inferential Comprehension	1320	1420	1420		
	Study Skills	1320	1470	1470		
1988	Total Reading	1300	1380	1380	1340 - 1400	
	Vocabulary	1300	1460	1460		
	Literal Comprehension	1300	1380	1380		
	Inferential Comprehension	1300	1360	1360		
	Study Skills	1300	1400	1400		

II. Developing a School Profile of Achievement

Although comparisons are a good way to start the process of using the assessment results effectively, they must be examined in terms of your own students and your own curriculum. The fact that a school has performed well in comparison with the state or district score gives very little useful information in itself. Performance must be measured in terms of the school's own goals and objectives.

Achievement in Content Areas.

Scores for the subcategories within each subject area are reported on pages 7 through 13 of the report. The number of items used is also listed. Examples of the actual items are given in the *Description of Test Content and Reporting Categories*, which should be in your school but is also available from the Department of Education.

Within most subject areas, the assessment material has been classified according to content and the type of thinking processes required. In the area of Reading, the categories reflect components of reading and the kind of passage on which the questions were based. Whatever your total score, it is important to start your evaluation by examining the individual categories.

Massachusetts Educational Assessment Program

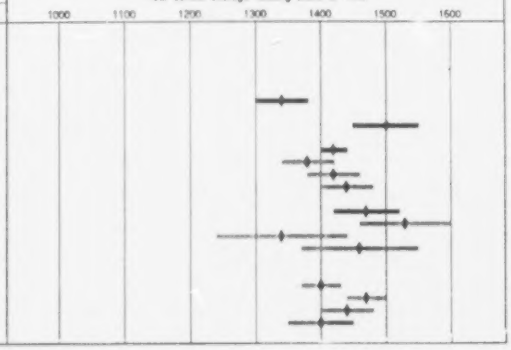
CURRICULUM AREA RESULTS FOR READING

ID:
District
School
Grade 8 Date April, 1990

The scaled scores for this school are presented on this page and pages 9, 11, and 13. The scaled score for the school and the number of questions that were included in the assessment are provided for each objective measured by the assessment. On the graphic display, the scaled score is shown as a diamond and the measurement error associated with that score as a bAr to both sides of the diamond. For example, your scaled score of 1340 in Vocabulary is shown as a diamond. When measurement error is taken into account, your Vocabulary score should be interpreted as being in the range of 1300 to 1380.

When the measurement error bars for two content areas overlap, the differences between them should be considered not statistically significant. When two diamonds are within 50 points of each other, the differences between them should be considered educationally significant.

Reporting Category	Number of Questions	Average Test Score: State	District	School
READING				
CONTENT AREAS				
I Vocabulary	21	1320	1340	1340
II Literal Comprehension	22	1300	1500	1500
III Inferential Comprehension	125	1320	1420	1420
A External perspective	39	1310	1380	1380
B Internal perspective - Analyzing Text	45	1320	1420	1420
C Internal Perspective - Evaluating Ideas	41	1320	1440	1440
IV Study Skills	45	1320	1470	1470
A Using Reference Materials	17	1320	1530	1530
B Following Directions	18	1310	1340	1340
C Summarizing/Organizing	10	1320	1460	1460
PASSAGE TYPES				
I Literary	77	1320	1400	1400
II Content	65	1320	1470	1470
III Practical	40	1310	1440	1440
IV non-passage related	21	1320	1400	1400



In the visual display, the diamond indicates the score which your students received, while the bar indicates the relative precision of that score. You will notice that some bars are longer than others. This is a function of three factors: the number of students in your school who took the tests, the number of questions in that category, and the percentage of students answering each item correctly. It is an acknowledgment of the fact that scores achieved on a test are only estimates of performance and that some scores are more precise than others. As a consequence, the general rule for inferring strengths or weaknesses is to pay particular attention only to those sub-categories where the bars do not overlap.

In interpreting these scores, it is important to remember that they do not represent an absolute value, but their value is determined by the scores that were achieved over the state as a whole in 1988. Practically, it is possible to achieve a comparatively low scaled score in a low-level skill, such as literal comprehension, and a high scaled score in a higher-level skill, such as inferential comprehension. This does not necessarily mean that the students in your school cannot identify material that is contained in the text. It merely indicates that they performed less well than others around the state in that skill. Subcategories that were answered well or poorly by most of the students, thereby showing little variance, are less reliable indicators of relative performance than subcategories in which there was a wide range of scores. To order to help you make a better judgment on relative strengths and weaknesses of your students, we have reported the mean and standard deviation of the raw scores for each category in Appendix II.

Achievement in terms of proficiency level

Scaled scores were developed to make accurate and valid comparisons; however, they are of little use in describing what students can and cannot do and can actually foster incorrect interpretation. For example, relative to the state as a whole, students in a school may receive a high score in chemistry while, at the same time, answering few of the chemistry questions correctly. Although they may be **perceived** to perform well in chemistry, their high score is a function, not of their own achievement, but of other students' **lack** of achievement. In order to relate the assessment results to actual achievement, proficiency levels have been included in school and district reports. They are introduced to add some meaning to the scaled scores, which have no inherent meaning of their own.

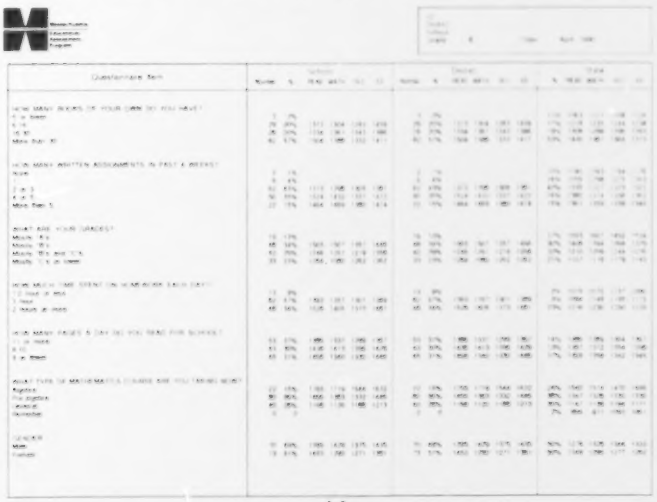
Proficiency Level	Students Scoring at This Level or Above				
	State	District		School	
	Percent	Number	Percent	Number	Percent
950	80%	122	85%	122	85%
1250	55%	96	67%	96	67%
1550	29%	52	36%	52	36%
1850	12%	28	20%	28	20%

Four levels of proficiency are defined. Students at Level I had scaled scores of 950; at Level II, 1250; at Level III, 1550; and at Level IV, 1850. (While school scaled scores can range only between 1000 and 1600, student scores are more variable and can exceed this range.) Descriptions for each level were developed by committees of Massachusetts teachers who were presented with sets of test questions that were answered correctly by a substantial majority of students at a given level, but were incorrectly answered by students at the next lower level. The committees reviewed these sets of questions to identify and describe the common skills represented by each set. The results of their analysis is summarized in Appendix I of this booklet. A more detailed description of the content knowledge and skills for each subject area is presented in a separate publication, *Description of the MEAP Proficiency Scales*.

The table included in the reports provides the number and percentage of students in your school who scored at each level and above. For example, the percentage of students at Level I includes all students in your school who have attained that level or above. As is evident from the table, the majority of students in the state have attained the skills and knowledge typical of Level II, whereas relatively few are at Level IV.

Student questionnaire results, including course work

The results of the student questionnaires appear on pages 14 and 15 of the reports. At the fourth and eighth grade levels, students were asked about their personal activities, such as the extent to which they read for pleasure, and about school practices that have been associated with achievement, such as the extent and type of homework assignments they completed. When all students replied to a question (as in the case of the fourth grade questionnaire and the "common questions" at grade eight), results are supplied for the entire school. In the case of questions that were answered by only a fraction of the students, the percentage and number of students answering are given. Although the attainment of those particular students is not given, it may be inferred from the statewide results.



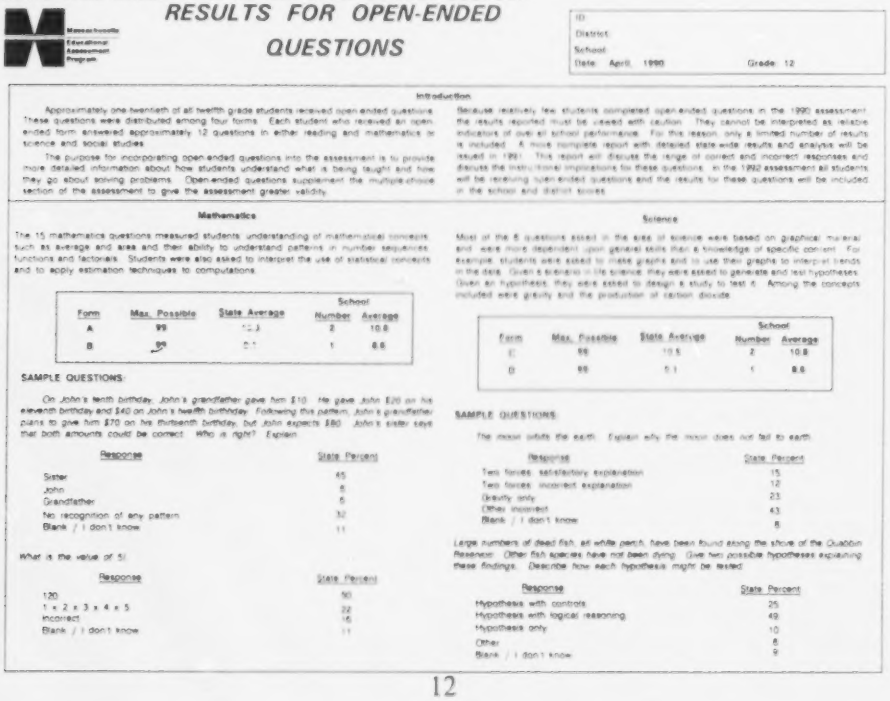
The questions that are contained in the report are those that are known to be related to achievement. Although it would be invalid to make the claim that the activities represented **cause** high achievement, they do indicate some of the factors that are associated with it. For each school, they may be helpful in identifying practices that should be encouraged among students and staff.

In the twelfth grade reports, scores are also given for students in different courses and for students who report different average grades. Students' "opportunity to learn" is a well-recognized factor in achievement. Statewide, the difference in achievement among students in different courses is dramatic. A study of this section will give you the opportunity to see if a disproportionately large number of students in your school are taking lower level courses. It will also provide evidence of grade inflation.

Finally, on page 14 the results of the local option question are included. Since the questions differ from district to district, no statewide results are given. To determine what the questions were for your district, you may need to contact the person within your district responsible for the test administration.

Open-ended results

The results for open-ended questions being reported for the first time this year. This is in anticipation of a change in the testing program in which all students will receive some open-ended questions as part of the assessment. From 1992 onward, the results of these questions will contribute to the total school scores.



Massachusetts Educational Assessment Program

RESULTS FOR OPEN-ENDED QUESTIONS

ID:
District:
School:
State: April, 1990 Grade: 12

Introduction

Approximately one-twentieth of all twelfth grade students received open-ended questions. These questions were distributed among four forms. Each student who received an open-ended form answered approximately 12 questions in either reading and mathematics or science and social studies.

The purpose for incorporating open-ended questions into the assessment is to provide more detailed information about how students understand what is being taught and how they go about solving problems. Open-ended questions supplement the multiple-choice section of the assessment to give the assessment greater validity.

Because relatively few students completed open-ended questions in the 1990 assessment, the results reported must be viewed with caution. They cannot be interpreted as reliable indicators of overall school performance. For this reason, only a limited number of results is included. A more complete report with detailed state-wide results and analysis will be issued in 1991. This report will discuss the range of correct and incorrect responses and discuss the instructional implications for these questions. In the 1992 assessment all students will be receiving open-ended questions and the results for these questions will be included in the school and district scores.

Mathematics

The 15 mathematics questions measured students' understanding of mathematical concepts such as average and area and their ability to understand patterns in number sequences, functions and factorials. Students were also asked to interpret the use of statistical concepts and to apply estimation techniques to computations.

Form	Max. Possible	State Average	School Number	School Average
A	99	12.3	2	10.0
B	99	9.1	1	8.0

SAMPLE QUESTIONS:

On John's tenth birthday, John's grandfather gave him $10. He gave John $20 on his eleventh birthday and $40 on John's twelfth birthday. Following this pattern, John's grandfather plans to give him $70 on his thirteenth birthday, but John expects $80. John's sister says that both amounts could be correct. Who is right? Explain.

Response	State Percent
Sister	45
John	6
Grandfather	6
No recognition of any pattern	32
Blank / I don't know	11

What is the value of 5!

Response	State Percent
120	50
1 x 2 x 3 x 4 x 5	22
Incorrect	16
Blank / I don't know	11

Science

Most of the 6 questions asked in the area of science were based on graphical material and were more dependent upon general skills than a knowledge of specific content. For example, students were asked to make graphs and to use their graphs to interpret trends in the data. Given a scenario in life science, they were asked to generate and test hypotheses. Given an hypothesis, they were asked to design a study to test it. Among the concepts included were gravity and the production of carbon dioxide.

Form	Max. Possible	State Average	School Number	School Average
C	99	10.5	2	10.0
D	99	9.1	1	8.0

SAMPLE QUESTIONS:

The moon orbits the earth. Explain why the moon does not fall to earth.

Response	State Percent
Two forces, satisfactory explanation	15
Two forces, incorrect explanation	12
Gravity only	23
Other, incorrect	43
Blank / I don't know	8

Large numbers of dead fish, all white perch, have been found along the shore of the Quabbin Reservoir. Other fish species have not been dying. Give two possible hypotheses explaining these findings. Describe how each hypothesis might be tested.

Response	State Percent
Hypothesis with controls	25
Hypothesis with logical reasoning	49
Hypothesis only	10
Other	8
Blank / I don't know	9

Approximately one-twelfth of fourth graders, one-sixteenth of eighth graders, and one-twentieth of twelfth graders received an open-ended form of the assessment. Since there were several forms, different students completed different questions. In any one school, very few or no students might have answered a specific open-ended question. For this reason, students' responses for each question are not given, merely their average score on the form which they received.

Unlike the multiple-choice questions, responses to the open-ended questions were not marked merely right or wrong but were classified according to a set of characteristics. A qualitative scale was used to evaluate responses to some questions. Others were classified according to type of error or correct responses. Responses of all students have been categorized and reported. A more complete discussion of the statewide results, along with the questions themselves will be reported separately in a series of publications entitled *On Their Own: Student Responses to Open-Ended Questions*. These reports will be available in the spring of 1991.

Although the information on the open-ended questions is extremely limited in the current report, it gives some indication of how well your students performed when asked to generate a response rather than select one. A more complete indicator can easily be obtained by administering the questions to all your students. Other questions, with statewide results, have been reported in the 1989 edition of *On Their Own*, which is available from the Department of Education.

CHAPTER FOUR

DEVELOPING AN ACTION PLAN

It is not the purpose of this handbook to tell you how to improve curriculum and instruction. There are other, better sources for this. Instead, its purpose is to help you use the assessment reports to look at your curriculum and instruction and to ask the right questions. The answers to these questions must come from you, with the help of your staff and other interested individuals.

In setting priorities, it is important to review your entire program. Although comparisons with state and district give relative performance, in no sense do they indicate how well your students have performed in any absolute sense. This is because the assessment is not a competency test; it does not put forth specific objectives which we believe that all students should master. Unlike Basic Skills testing, there is no absolute standard of knowledge involved. As a corollary to this, a relatively high total score on the assessment does not indicate a uniformly superior program. In order to judge how your program can be improved, it is necessary to look at the specific curriculum area results for indications of strengths and weaknesses.

Developing An Action Plan

What are the strengths and weaknesses of our instructional program?
Compare specific objectives with overall content score

↓

Are we satisfied that this is an area of weakness?
Compare scores with those from other sources

↓

Is this topic or skill part of our intended curriculum?
Review written curriculum, textbooks, teaching materials

↓

Is this topic or skill part of the actual curriculum?
Use teacher questionnaires as focus for discussion

↓

What are the other factors affecting performance?
- Use teacher/student questionnaires as focus for discussion
- Study state report

↓

- Form committees
- Set time lines
- Give progress reports

Question 1: What are the strengths and weaknesses of our instructional program?

Scores for the subcategories within each subject area are reported on pages 7, 9, 11, and 13 of the report. The number of items used is also listed. Examples of the actual items are given in the *Description of Test Content and Reporting Categories*, which should be in your school but is also available from the Department of Education.

The first step in evaluating your program is to examine your performance on the various components of the assessment. Within most subject areas, the assessment material has been classified according to content and to the type of thinking required. In the area of Reading, the categories reflect components of reading and the kind of passage on which the questions were based. Whatever your total score, it is important to start your evaluation by examining the individual categories.

The simplest way to identify areas of weakness in each subject area is to use the total score as a reference and draw a vertical line from that point. You can then note how the bar for each subskill lies in relation to that line. Bars that fall to the right of the line represent relative strengths, while those that fall to the left signify relative weaknesses.

Question 2: Are we satisfied that this is an area of weakness?

- *Do other test scores verify the assessment scores?*
- *Do teacher judgments confirm results?*
- *Is there other evidence to confirm results?*

Whether a subcategory can be considered an area of weakness is a matter of judgment. That judgment rests on a number of factors which should be considered by your curriculum area experts and the teachers involved.

For most schools the assessment results are one of many sources of information on school performance. Commercial standardized test results are available, as are scores on department-wide tests and classroom measures. The extent to which information obtained from these other sources converges is a measure of its validity.

On the other hand, the assessment may point to weaknesses that exist but are not measured by other instruments. The assessment was designed to cover a broad range of knowledge and types of thinking that are representative of each subject area. As a result it measures, along with more traditional skills, certain strategies and kinds of thinking that, although considered important by content area experts, are not explicitly

taught or tested in some schools. For example, the metacognitive skills that are included in our reading framework are a departure from standardized reading tests. Although this particular innovation was designed to make the assessment results more relevant to instruction and more in keeping with current research on the reading process, achievement in this area may not be congruent with scores obtained on more traditional reading tests. This lack of congruence is an important consideration in interpreting test results. Unless a test is aligned with the instructional program, it will not be sensitive to effects of that program.

Curriculum evaluation is not a mechanical process that can be followed through a series of check-lists. Basically, it is a matter of examining the evidence and making informed judgments. For this reason, when choosing an area for curriculum improvement, it is important to examine the various sources of information carefully before deciding where resources should be placed.

Question 3: Is this topic or skill part of our intended curriculum?

- *Does this topic or skill form part of our written objectives?*
- *Is it covered adequately in the textbooks and curriculum guides?*
- *Is instruction supported by curriculum materials and other learning aids?*

Recently, educators have defined three different aspects of the word "curriculum." They refer to the written curriculum—that which is contained in the goals, objectives, and scope and sequence charts in our schools—as the **intended curriculum**. They note that this intended curriculum is distinct from the **actual curriculum**—that which is filtered through the teacher and is taught in the classroom itself. Finally, there is the **achieved curriculum**, what is learned by students and usually reflected in test results. In order to understand why a particular curriculum has or has not been achieved, it is necessary to examine how the first two terms are understood and employed.

Although a written curriculum exists in many schools, the assessment offers an opportunity for school staff to come together to review and revise, if necessary. Using the descriptions contained in the *Description of Test Content and Reporting Categories* and the appropriate sample items as guides, the committee should consider the material covered in their school(s) through their curriculum objectives and instructional materials. The subcategories in each subject area represent different types of knowledge, skills, or kinds of thinking. Committees should look at these categories and ask:

> *Does this content or type of thinking match our current list of curriculum objectives? Should it? If so, by what grade level? For all students?*

The intended curriculum is not only reflected in a set of written objectives. To be meaningful, it must also be supported by textbooks, curriculum materials, and other learning aids. At this stage, when examining the intended curriculum, it is important that committees go beyond what it written to what actually exists for teachers to use and how it is interpreted in learning materials.

Each school must decide for itself what aspects of the subject are most relevant and beneficial for its students. This will depend not only upon the characteristics and previous experiences of the students, but also upon the particular emphases of your curriculum and the resources available. Remember that the assessment was designed to include all the skills and concepts that are considered appropriate in the subject area tested. Given the broad scope of the assessment, it is highly unlikely that any school would be able to cover thoroughly every subcategory, nor would such an endeavor be productive given the time available.

On the other hand, it is important to consider your relative performance on the different subskills. Very low scores can be flags for attention. Since the assessment includes very easy as well as very difficult items within each subskill, a low score suggests that your students may be missing out on fundamental concepts within the subject area as a whole.

Question 4: Is this topic or skill a part of the actual curriculum?

- *Is this skill or topic actually being taught in the classroom? To what level of mastery?*
- *Are there adequate resources?*
- *Are teachers interested and supportive of this curriculum?*
- *Is there the support for teachers, including in-service?*

If the skill involved is one that has formed part of your intended curriculum and remains an area of weakness, it is important to examine the actual or delivered curriculum. In other words, is the skill being taught in the classroom and how is it being taught?

There are many reasons why content areas or skills are not taught or are taught sketchily. These include lack of time, lack of supporting materials, lack of interest on the part of the teacher or students, and lack of confidence by the teacher that s/he can teach it effectively. Each reason carries its own solution. For example, teachers may believe that the topic is inappropriate for their students or a portion of their students. They may believe that the topic is untimely, or presupposes experiences that children have not yet attained. On the other hand, the skill or topic may not be taught because of lack of resources. Books or appropriate materials may not be available. Teachers themselves may not have sufficient training to feel competent to teach it to the degree that ensures learning. To acknowledge

why a skill or topic is not being taught is as important as discovering whether it is being taught.

The teachers' questionnaire should be a valuable source for promoting discussion. We have encouraged schools to keep those questionnaires (returning only the answer sheets) so that they could be used, with the state summary report, to structure discussion on instructional methods and the actual content of instruction. At the fourth and eighth grade levels, these questionnaires contain actual items which were used on the assessment. It will be useful to confirm if these items reflect the kind of knowledge that you expect your students to know and understand and to what extent consensus exists among teachers of different classes.

In a small setting in which thoughts and feelings can be expressed freely, it may be possible for the group to discuss each of the objectives in turn. In other settings, it may be more useful to conduct a simple written survey, asking each teacher to indicate to what extent (mastery, introductory, or reviewed) each of the objectives is taught in each of his or her classes. In any event, keep your discussions frank and nonjudgmental. Do not set out with the expectation that every subskill in the assessment (or even in your curriculum guide) can be taught to mastery by every teacher. Instead, keep track of the reasons why objectives are not being taught. Consider them as problems to be solved, not as "excuses" for a job not done.

Question 5: What are the other factors affecting performance?

If the objective is in both the intended and the actual curriculum, but students are performing poorly in relation to their performance in the subject as a whole, you must ask another set of questions.

- *Do our students understand the concepts as well as we believe they do?*

The assessment was designed to reach beyond factual knowledge and to tap students' understanding of concepts and their ability to apply those concepts to familiar and unfamiliar situations. In many instances, the items differ from those questions found in textbooks at the end of chapters and neatly tie in with the previous instruction. Students who are adept at answering the latter are often unprepared for more challenging questions in different contexts. The *State Summary Report* discusses some of the discrepancies that occur between students' knowledge of facts and their performance when asked to use those facts in solving problems.

Classroom discussion can also give false information when teachers consistently ask for correct responses. For example, it has been observed that students' lack of understanding of mathematical and scientific terms is often masked by their ability to supply the correct term when asked. Furthermore, in classroom discussions, if teachers predominantly call on volunteers who know the answer, they may get an inflated impression of how the class is

doing as a whole. Some close questioning of a wider group of students may give a more representative assessment of progress.

At the eighth and twelfth grade levels, students were asked to give a general indication of the grades that they receive at school. The assessment scores for students responding to each category are given for school (when the number of students is greater than 20), district, and state. This comparison will give you some indication of possible grade inflation (or deflation) among the teachers in your own school. Although some students may exaggerate their grades, it has been found that they are generally quite accurate in their reports, and the overall average response is a reliable indication of grade achievement.

In addition, the results from the open-ended sections of the assessment may suggest ways teachers might use open-ended questions to assess their students' understanding of different topics. Copies of the reports on these results will be sent to all schools early in 1991.

- *Are our expectations too low?*

Students may be capable of more than we demand of them. For example, studies have shown (and our results have confirmed) the strong influence of prior knowledge in many "thinking" tasks. When students have knowledge of the content area, they are able to interpret visual data and to answer comprehension questions that they find difficult in other contexts. In other cases, students often use informal knowledge to respond correctly to problems in areas not covered in the curriculum. This suggests that expectations of students' performance may limit what we allow them to learn.

- *Is learning articulated throughout the grades and reinforced across subject areas?*

Although the assessment takes place at a particular grade level, it measures all the learning that has taken place up to that time. Students may forget unconnected facts, but we would expect that their basic understanding of the subject would continue to be applied to more recent instruction. On the other hand, one of the most important findings of the assessment is the lack of reinforcement of learning across subject areas, despite general themes that permeated all subjects. The extent to which subjects are interrelated throughout the curriculum is discussed in the *State Summary Report* where specific examples are given. Schools should study this report and examine their own instruction for evidence that writing is not isolated to language arts, graphic interpretation is not the sole province of the science teacher, and reading is taught across content areas.

- *Are we attempting to reach beyond school gates to involve parents and others in promoting our objectives?*

The student questionnaire results give evidence for the strong association between out-of-school activities and performance on the assessment. Average assessment scores are given at state, district, and school levels for those students who responded in different ways to questions regarding free time and work activities. The percentage of students in your school who responded to the various options is also given to help you compare the activities of your students with others.

Despite the fact that the school has a limited role in affecting student behavior outside of its walls, it can encourage the reinforcement of learning at home. The strong effect of social factors such as parental education may be due, in large measure, to this kind of reinforcement. For example, the statewide summary report shows that certain student activities are associated with performance. That these activities (e.g., reading for pleasure, television watching, etc.) are related to achievement does not mean that they cause achievement. Nevertheless, students who spend their time outside of school applying their instruction in different contexts are essentially extending and reinforcing what they have learned during the school day.

In addition, schools have the opportunity of bringing the outside world into the school. Class trips, activities based on practical and meaningful problems, guest speakers—all contribute to understanding the relevance of school learning to everyday life.

Final Steps: Preparing an action plan

- *Does the plan document the most important weaknesses? strengths?*
- *Is the plan based on thorough diagnosis from credible evidence?*
- *Are steps/recommendations/plans for proceeding clearly specified?*
- *Are resources, participation, and time delineated? realistic?*
- *Is an evaluation plan for determining the effectiveness of the proposed plan detailed?*

If your plan is to be effective, there are some important points to remember. One is the need for enlisting the support of all involved. To the greatest extent possible, order your priorities on the basis of consensus. Unless there is general interest and enthusiasm for the effort, little will be accomplished. Secondly, it is important to be realistic in terms of the resources and human effort that will be involved. Before mounting any effort, review your textbooks and available materials to make sure that you have the resources to teach the skills and content that you have chosen. Be prepared to spend more time and energy than anticipated on instituting the new areas of instruction. Finally, don't concentrate on a single year. Skills

and knowledge are cumulative. It is important that curriculum objectives, however specific, be seen as the logical extension of earlier learning, to be reinforced in the later grades.

In a complex process such as described here, it is sometimes possible to lose the way. Groups become distracted; other commitments take precedence. The impact of the assessment results begins to recede. For these and other reasons, it is important to do two things: set a time line and plan a report.

Setting realistic deadlines for stages in your evaluation and for performance goals in your action plan will strengthen your commitment to the project. A general time line can be anticipated, with deadlines for specific steps such as teacher surveys or reviews of textbooks, to be implemented along the way. The fact that you have a sequential plan and regard it seriously will help participants feel a sense of progress.

Secondly, regular progress reports to faculty and parents will help focus attention on the overriding questions of "what have we learned in our evaluation?" and "what are we doing about it?" Reports should address perceptions and problems, as well as more tangible results. They will form an important public record of your progress in school improvement.

APPENDIX I

PROFICIENCY SCALES

GRADE FOUR

950

General characterization: Students at this level are beginning to master the basic skills and concepts of the elementary school curriculum. They perform well when questions are set in contexts that they recognize as part of their experience and when material is presented visually.

Reading: They are adept at reading simple, short passages that are supplemented by graphics. Although they can locate information within the text, their inferences are based upon common sense rather than upon material found in the passage. Their vocabulary is limited to simple, commonly-used words.

Mathematics: They have mastered basic arithmetic facts and operations, recognize the relative size of numbers, and can solve simple, one-step problems.

Science and social studies: In these areas they appear to rely more upon their own experiences and common sense than upon classroom instruction. Although they show familiarity with common scientific facts that are the typical focus of children's literature and television, they appear to have much more limited knowledge of the society about them.

1250

General characterization: Students at this level are adept at the skills that constitute the main focus of the elementary school curriculum. Within familiar contexts, they show an ability to reason logically.

Reading: Their reading material is no longer limited to simple short passages, but they are able to remember details and make simple inferences concerning narratives of some length and complexity.

Mathematics: They are adept in carrying out operations with whole numbers and simple decimals. They can translate simple word problems into number sentences and can solve multi-step word problems involving addition and subtraction. In addition, they can read and interpret frequency graphs and scales.

Science and Social Studies: They exhibit an understanding of the more fundamental ideas of the subject areas, such as roles and social groups in social studies, animal classification, and states of matter in science. They also have some familiarity with basic research concepts in both areas.

1550

General characterization: Students at this level are keen observers, paying attention to details. At the same time, they are beginning to think in terms of larger systems. This ability to recognize abstractions allows them the confidence and flexibility to reason through complex word problems, analyze content, and distinguish between necessary and irrelevant detail.

Reading material: Their reading ability is such that they can view passages in their totality, rather than as a sequence of events. When presented with complex and detailed material, they do not lose their way, but can evaluate information and make judgments, distinguishing the best response among a set of plausible ones.

Mathematics: These students are beginning to view number in its own right, not solely as the property of things. They recognize number sequences, patterns, and simple functions and are beginning to have a sense of variables.

Science: Although these students have a much broader knowledge level than students at the 1250 level, they can go beyond mere scientific facts to show an understanding of such fundamental concepts as the function of organs and the effect of gravity.

Social Studies: Showing a greater awareness of community, levels of government, and cultural diversity, these students also appear to have a sense of the past, recognizing the sequence of historical events.

1850

General characterization: Students at this level have acquired a broad knowledge that goes beyond the taught curriculum and show an ability to use that knowledge as tools for logical analysis and reasoning.

Reading: They are adept readers, with a large vocabulary and consciousness of themselves as readers. They can read a wide range of fiction, factual material, and simple poetry, and are beginning to recognize such formal aspects of language as poetic conventions and imagery.

Mathematics: These students can reasoning numerically. They recognize simple mathematical symbols and show a knowledge of basic statistical concepts.

Science: They show a deeper understanding of underlying concepts and relationships. They are able to analyze, reason, evaluate, and synthesize.

Social Studies: They show clear reasoning ability: inferring from data, recognizing assumptions, and justifying their responses. Their knowledge of history goes beyond the recognition of benchmark dates to indicate an understanding of some underlying historical currents, as well as a sense of time.

GRADE 8

950

General characterization: Students at this level are have mastered some of the skills and concepts of the elementary school curriculum; however, their understanding appears to be limited to the specific and practical. They perform best when questions are set in contexts that they recognize as part of their own experience.

Reading: Their level of reading is characterized by high-interest informational material. Although they can locate information within these texts, their inferences are based upon common sense rather than upon material found in the passages.

Mathematics: These students are proficient with operations of whole numbers and can translate simple word problems into number sentences. They are adept at visual tasks—interpreting frequency graphs and recognizing common geometric shapes.

Science and Social Studies: They are knowledgeable about those aspects of the subjects that apply to their personal lives, i.e., nutrition in science and personal rights in social studies.

1250

General characterization: These students have mastered much of the knowledge and skills that are typical of the taught curriculum. They are critical readers, evaluating material and recognizing assumptions. They are also adept at using visual material such as interpreting more advanced graphs, visualizing figures, and abstracting information from cartoons.

Reading: They are active readers of a wide range of both narrative and informational passages, making inferences about character and motivation and identifying assumptions. They recognize of some of the formal aspects of reading, such as the purpose of different typographical conventions (e.g.,italics).

Mathematics: They show confidence in applying operations with decimals and fractions and can solve multi-step word problems, including those with integers. They recognize variables, can evaluate expressions, and use formula.

Social Studies: These students possess a general knowledge of the world about them, as well as knowledge of history and government. They also show an understanding of some basic social studies concepts, such as the effects of change.

Science: They can go beyond the everyday aspects of science to recognize instances of basic scientific concepts (e.g., evaporation and condensation) and to related individual elements within a system. They are also familiar with scientific convention and survey techniques.

1550

General characterization: These students think in terms of systems rather than discrete components. At the same time, they show a great attention to detail and a persistence in working through material. In addition, they are able to evaluate possible solutions by discriminating among nuances of meaning. They go beyond the taught curriculum to relate information from many sources.

Reading: These students are fluent readers and can handle difficult vocabulary. They appreciate the purpose and function of writing conventions and author's voice.

Mathematics: In addition to a thorough knowledge of numerical operations and understanding of basic numerical concepts (decimals, square roots) that are part of the eight grade curriculum, they can reason numerically.

Science: These students show a deeper understanding of scientific concepts by relating those concepts to a variety of situations that are different from the contexts in which they are taught. They can draw conclusions based on careful analysis of data and can evaluate experimental designs.

Social Studies: These students go beyond the facts and taught definitions of the curriculum to recognize the underlying concepts. They can interpret and integrate information from different sources, including unfamiliar graphs and charts.

1850

General characterization: These students not only possess a broad base of detailed information about the subject areas tested, they are distinguished by their ability to analyze material and reason from it. Furthermore, in all areas, they possess an extensive vocabulary of technical terms that are appropriate to the subject.

Reading: These students can make fine discriminations in meaning. They recognize nuances in meaning and tone of voice and have a sophisticated understanding of craft of writing — poetic forms, irony, and stylistic devices.

Mathematics: These students understand of basic algorithms and problem solving strategies. Unlike students at the previous level, they are adept at using proportional reasoning in their analysis of problems.

Social Studies: These students possess detailed historical information. They are familiar with technical terms, such as scapegoating and lobbying. They are critical readers, recognizing assumptions and biases, as well as being thoughtful writers, synthesizing information and ideas from a number of different contexts.

Science: These students possess a great store of detailed science information, which they can synthesize when presented with a problem. They can analyze systems and discover relationships.

GRADE 12

950

General characterization: Students at this level continue to perform best in a personal and practical context. They have acquired a set of academic skills that are useful in everyday life.

Reading: Their reading material consists of high-interest informational material and narrative passages that deal with human relationships. They can locate information within these texts, interpreting words in context and recognizing some figurative language.

Mathematics: In addition to the basic operations, they can add and subtract integers and can use coordinates to read charts. They can recognize variables and evaluate expressions.

Social Studies: Having accumulated a store of knowledge about the history and structure of society, they are able to use it in reasoning about real-life situations. They are also able to read and interpret graphical information, and possess a knowledge of place geography.

Science: They are familiar with basic scientific concepts, such as forms of matter and the elements of a food web, as well as everyday aspects of health.

1250

General Characteristic: Students at this level show a greater maturity and a broader base of knowledge than eight graders at the 1250 level; however, many of the characteristics of their thinking and some of their skills are comparable. In general, they are critical readers of text, but are less adept at integration and abstraction.

Reading: They show an understanding of a wide and sometimes stylistically difficult range of passages in history, sociology, and current events. They are able to summarize the main points of the articles, recognize assumptions, and evaluate the material presented. They can make inferences about character and motivation. They can actively construct meaning and recognize authors' assumptions. In addition, they can recognize the purpose of different typographical conventions (e.g.,italics).

Mathematics: These students are adept at basic operations with whole numbers, fractions, and integers, and can and perform basic algebraic manipulations, but appear to have little understanding of more abstract concepts or mathematical reasoning.

Science: They have a broad based knowledge of concepts in the area of life science and can answer questions that are typical of work in biology beyond the basic level. They do not show the same breadth of knowledge in the physical sciences. They also appear to have a fundamental understanding of the nature of science.

Social Studies: Their broader knowledge base and more proficient reading level allows them to display the kind of critical, active reading that is important to the subject. They show also a knowledge of current events and factual modern history, as well as an understanding of fundamental civic concepts such as representative democracy.

1550

General Characteristics: These students are primarily distinguished by their use of conceptual and logical analysis to solve problems. They also show a great attention to detail and can integrate information from the text itself, from other reading, and from their own experience in their construction of meaning. They possess a broad base of academic knowledge, both in history and in all major areas of science, and they are able to integrate and apply this knowledge to problem situations.

Reading: These students are adept readers of detailed technical material. At the same time, in their reading of more general literature, they show an appreciation of figurative language and the use of author's voice. Their vocabulary is extensive, including both sociological and economic terms, as well as figurative terminology.

Mathematics: These students understand mathematical concepts, such as scientific notation and integers, and are adept at using both geometric knowledge or algebraic manipulation to solve problems. They show evidence of an algebra curriculum.

Science: Students at this level display acknowledge of basic information in the physics, chemistry, and earth science curriculum. On the other hand, what is most evident about these students is their ability to analyze questions at a fine level of detail, to evaluate, and to apply reasoning skills.

Social Studies: These students show a broad knowledge of current events, civics, and modern history, as well as a knowledge of ancient civilizations. More important, they are able to integrate this information in response to specific situations.

1850

General Characteristic: These students use a broad and detailed base of knowledge and understanding to solve problems and to construct meaning. They appear to have mastered all the elements of the secondary school curriculum in the subjects tested. They recognize the meaning of specialized terms in economics, as well as in other areas. Most striking is the independence in thinking.

Reading: They are thoughtful readers, recognizing and rejecting stereotypical conclusions. In addition, they are able to distance themselves from the meaning of the text to take into account the craft of writing itself.

Mathematics: These students possess great computational fluency, as well as evidence of understanding of more advanced mathematical topics. They show an unusual flexibility in problem solving, using divergent thinking strategies to tackle unfamiliar problems.

Science: The ability to grasp more complex scientific principles and to apply them in experimental and practical contexts is typical of these students. They exhibit an understanding of the scientific process by applying the principles of science in both experimental and practical contexts.

Social Studies: Students at this level are primarily distinguished by their detailed knowledge of history and governmental processes. They are able to abstract information from different sources and apply it to unfamiliar contexts.

APPENDIX II

1990 MASSACHUSETTS EDUCATIONAL ASSESSMENT PROGRAM MEAN PERCENT CORRECT BY OBJECTIVE

GRADE 4

Objective	Mean	Standard Deviation
READING	70.32	8.07
Vocabulary	69.67	11.56
Literal Comprehension	82.52	8.18
Inferential Comprehension	66.69	9.51
Analyzing Text	68.99	8.95
Evaluating Ideas	66.37	9.17
Study Skills	74.60	10.14
Literary	67.75	8.99
Content	70.51	8.64
Practical	71.48	8.51
MATHEMATICS	61.56	8.44
Numeration	67.56	9.84
Number Theory	67.24	10.72
Whole Numbers (Operations)	69.41	9.69
Fractions and Decimals	63.00	11.03
Properties of Operations	53.41	16.02
Variables and Relationships	61.87	10.09
Using Instruments	64.39	9.98
Units (of Measurement)	55.93	10.44
Perimeter, Area, Volume	42.52	14.14
Plane/Solid Figures	55.60	12.60
Transformations/Spatial Visualization	60.17	10.44
Estimation and Reasonableness	55.38	10.70
(Problem Solving) Strategies	59.13	10.65
Relevant Information	61.85	11.96
Probability and Statistics	47.75	11.75
Graphs, Tables, Charts	67.33	10.75

Conceptual Understanding	61.40	8.71
Procedural Knowledge	63.06	8.86
Problem Solving	60.99	9.26
Reasoning & Analysis	61.90	10.42
SCIENCE	64.11	7.62
Designing Investigations	68.87	10.26
Data Gathering/Reduction	68.29	9.67
Data Interpretation	72.83	9.92
Nature of Science	60.35	14.05
Characteristics of Life	62.03	9.10
Health	64.87	9.60
Animal Life	72.58	10.14
Plant Life	62.05	11.06
Ecology and Environment	56.68	10.08
Astronomy	61.68	10.45
Meteorology	59.58	9.66
Geology/Nat.Res./Oceanography	61.80	10.25
Matter	62.35	10.31
Energy	59.57	10.32
Force and Motion	63.36	9.09
Knowledge/Comprehension	65.72	7.75
Application/Higher Order	62.34	7.98
Scientific	64.41	7.74
Personal	66.71	8.13
Technological, Societal, Historical	58.95	9.48
SOCIAL STUDIES	69.57	8.12
Specifics of History	68.70	8.58
Sense of History	71.39	10.37
Government	67.79	11.10
Citizenship	78.95	8.66
Physical Geography	69.77	11.53
Human Geography	68.10	11.58
Economic Environment	70.16	10.51
Social Environment	73.06	10.11

Multicultural Environment	68.42	9.20
Map Skills	81.53	8.24
Graphic Representation	73.36	9.80
Research Skills	69.16	10.65
Analysis	60.04	9.94
Evaluation	53.23	10.69
Knowledge/Comprehension	72.50	8.09
Application/Higher Order	67.46	8.44

GRADE 8

Objective	Mean	Standard Deviation
READING	68.05	7.41
Vocabulary	68.48	9.95
Literal Comprehension	74.89	7.65
Inferential Comprehension	61.98	9.06
Analyzing Text	73.41	7.51
Evaluating Ideas	64.54	8.16
Using Reference Materials	66.02	8.96
Following Directions	65.77	9.13
Summarizing/Organizing	71.07	9.70
Literary	68.13	7.34
Content	65.81	8.48
Practical	68.38	8.09
Non-Passage Related	74.00	9.04
MATHEMATICS	57.66	9.15
Numeration	66.46	12.01
Number Theory	54.53	10.14
Whole Numbers (Operations)	58.94	9.85
Fractions	57.62	11.12
Decimals	62.68	11.33
Percent	43.06	11.37
Integers	59.51	14.77
Properties of Operations	50.59	12.09

Variables and Relationships	54.01	10.99
Using Instruments	62.12	10.88
Units (of Measurement)	57.58	9.77
Perimeter, Area, Volume	47.16	12.35
Plane/Solid Figures	50.39	9.72
Transformations/Spatial Visualization	63.21	9.35
Estimation and Reasonableness	65.21	10.25
(Problem Solving) Strategies	59.97	10.11
Relevant Information	62.43	11.32
Probability and Statistics	64.27	10.50
Graphs, Tables, Charts	60.92	10.05
Conceptual Understanding	60.09	9.24
Procedural Knowledge	58.67	10.19
Problem Solving	54.92	9.26
Reasoning & Analysis	50.14	9.30
SCIENCE	54.49	6.81
Designing Investigations	60.57	7.90
Data Gathering/Reduction	56.17	8.90
Data Interpretation	53.56	9.92
Nature of Science	67.05	12.80
Characteristics of Life	50.39	8.02
Health	65.31	8.08
Animal Life	58.10	8.85
Plant Life	48.00	8.50
Ecology and Environment	58.06	8.07
Astronomy	61.18	8.89
Meteorology	45.86	9.74
Geology/Natural Res./Oceanography	48.35	8.77
Matter	54.91	9.43
Energy	48.63	7.70
Force and Motion	55.71	8.52
Knowledge/Comprehension	54.14	7.09
Application/Higher Order	54.93	6.82
Scientific	54.18	6.61
Personal	56.12	8.40
Technological, Societal, Historical	54.65	8.02

SOCIAL STUDIES	57.00	7.32
Pre-1865 U.S./American History	55.92	9.07
Post-1865 U.S./American History	58.10	8.79
Contemporary Scene (U.S.)	53.54	8.46
World History	49.54	8.04
Government	50.98	9.20
Citizenship	63.65	8.18
Physical Geography	57.99	8.94
Human Geography	57.03	10.48
Economic Environment	57.65	9.07
Social Environment	60.22	9.43
Multicultural Environment	62.84	9.62
Map Skills	55.32	9.24
Graphic Representation	60.57	10.01
Research Skills	62.90	10.21
Analysis	58.35	9.52
Evaluation	53.30	7.93
Knowledge/Comprehension	57.93	7.35
Application/Higher Order	56.19	7.56

GRADE 12

Objective	Mean	Standard Deviation
READING	70.60	7.80
Vocabulary	71.93	9.00
Literal Comprehension	78.16	7.45
Inferential Comprehension	67.06	7.81
Analyzing Text	70.43	8.28
Evaluating Ideas	68.98	8.27
Using Reference Materials	69.79	10.41
Following Directions	75.22	8.04
Summarizing/Organizing	75.11	8.56
Literary	63.97	7.13
Content	60.56	7.38
Practical	66.46	6.69
Non-Passage Related	76.86	8.49

MATHEMATICS	59.99	8.60
Numeration	65.61	10.33
Number Theory	59.38	10.07
Whole Numbers (Operations)	63.51	9.33
Fractions	62.97	9.57
Decimals	65.46	9.64
Percent	58.24	9.59
Integers	68.13	12.55
Properties of Operations	65.06	13.12
Algebraic Manipulations	58.69	13.90
Relations/Functions	54.68	9.75
Equations/Inequalities	54.07	12.09
Using Instruments	69.53	8.03
Units (of Measurement)	55.25	9.50
Perimeter, Area, Volume	51.59	11.04
Plane/Solid Figures	56.62	10.68
Transformations/Spatial Visualization	56.50	7.83
Estimation and Reasonableness	69.49	7.35
(Problem Solving) Strategies	58.95	9.36
Relevant Information	67.34	8.86
Probability	69.50	8.32
Graphs, Tables, Charts	65.77	8.27
Statistics	52.82	7.20
Conceptual Understanding	60.68	7.68
Procedural Knowledge	61.21	10.87
Problem Solving	58.80	8.24
Reasoning & Analysis	56.62	10.19
SCIENCE	56.08	6.67
Designing Investigations	67.32	8.91
Data Gathering/Reduction	58.33	8.78
Data Interpretation	53.58	8.44
Nature of Science	57.32	10.22
Characteristics of Life	55.19	8.78
Health	69.21	8.00
Animal Life	57.12	7.42
Plant Life	51.92	7.81
Ecology and Environment	53.01	7.63

Astronomy	58.32	8.39
Meteorology	57.20	8.74
Geology and Natural Resources	52.61	8.16
Oceanography	53.76	8.29
Matter	52.80	8.89
Energy	52.43	6.41
Force and Motion	57.77	6.88
Knowledge/Comprehension	56.26	6.45
Application/Higher Order	55.77	7.46
Scientific	55.43	6.61
Personal	61.70	7.34
Technological, Societal, Historical	56.12	7.58
SOCIAL STUDIES	60.40	7.33
Pre-1865 U.S./American History	61.78	8.48
Post-1865 U.S./American History	59.75	8.48
Contemporary Scene (U.S.)	63.22	8.84
Prehistoric and Classical Civilizations	52.45	7.87
Middle Ages (W./E.:500-1600 A.D.)	44.56	7.59
Modern World History(Renais.-1900)	47.16	8.11
20th Century World History	54.63	9.25
Principles of Government	60.88	9.93
Components of Government	61.94	9.31
Processes of Government	54.12	9.85
Citizenship	73.03	7.63
Physical Geography	63.15	9.41
Human Geography	63.49	9.12
Economic Environment	56.15	8.28
Social Environment	65.31	8.47
Multicultural Environment	67.28	8.36
Map Skills	68.65	8.73
Graphic Representation	68.52	8.88
Research Skills	57.77	10.08
Analysis	62.04	10.25
Evaluation	54.41	8.38
Knowledge/Comprehension	59.87	7.20
Application/Higher Order	61.04	7.72

MASSACHUSETTS DEPARTMENT OF EDUCATION

REGIONAL CENTERS

Paul Burnim
Greater Springfield Regional Center
Macek Drive
Chicopee, MA 01013
413/594-8511

Laurie Slobody
Central Mass. Regional Center
Beaman Street, Route 140
West Boylston, MA 01853
617/835-6266

Amy Sosman
Northeast Regional Center
Jefferson Office Park
790 Turnpike St.
North Andover, MA 01845
508/689-0164

Pat O'Brien
Southeast Regional Center
P.O. Box 29
Middleboro, MA 02346
617/947-1231

Related publications about the Massachusetts Educational Assessment Program:

- *The Massachusetts Educational Assessment Program: 1988 Statewide Summary*
- *Using the School Report*
- *Reading and Thinking: A New Framework for Comprehension*
- *Science in the Elementary Schools*
- *Moving Geometry From the Back of the Book: A Report on Geometry and Measurement in the 1986 Assessment*
- *Social Studies Report*
- *On Their Own: Student Response to Open-Ended Tests in Reading*
- *On Their Own: Student Response to Open-Ended Tests in Mathematics*
- *On Their Own: Student Response to Open-Ended Tests in Science*
- *On Their Own: Student Response to Open-Ended Tests in Social Studies*

Advanced Systems IN MEASUREMENT & EVALUATION, INC.
P.O. BOX 1217 • 171 WATSON ROAD, DOVER, NEW HAMPSHIRE 03820 • (603) 749-9102

BEST COPY AVAILABLE